Five-dimensional Gauge Theories and Unitary Matrix Models

Martijn Wijnholt

Department of Physics, Jadwin Hall, Princeton University

Princeton, NJ 08544, USA

Abstract

The matrix model computations of effective superpotential terms in $\mathcal{N} = 1$ supersymmetric gauge theories in four dimensions have been proposed to apply more generally to gauge theories in higher dimensions. We discuss aspects of five-dimensional gauge theory compactified on a circle, which leads to a unitary matrix model.

January 2004

1. Overview

In the physics of four dimensional $\mathcal{N} = 1$ supersymmetric gauge theories a special rôle is played by the F-terms. The low energy effective superpotential terms obey stringent constraints and can often be determined exactly. Pushing the simplification to its extreme, it was argued in [1] that for a large class of examples the computation can be done with only the planar diagrams in an auxiliary matrix model. The canonical example is that of an $\mathcal{N} = 2$ gauge theory deformed to $\mathcal{N} = 1$ by a superpotential, where the planar diagrams can be summed up exactly.

In [2] it was suggested that matrix model methods can also be used to analyse F-terms in higher dimensional supersymmetric gauge theories. Consider for instance a five dimensional $U(N)$ gauge theory with minimal $\mathcal{N} = 2$ SUSY. The matrix model methods of [1] apply to theories with only four supercharges and isolated vacua so we cannot tackle such a theory directly. However, suppose we put the theory on a circle of radius R and coordinate t. The resulting theory has a Coulomb branch parametrised by the eigenvalues of the complexified Wilson line

$$U = P \exp(\int dt\, A_t + \varphi). \tag{1.1}$$

Now one may turn on a superpotential $W(U)$ with isolated critical points. It also breaks the number of supercharges to four. This theory exhibits gluino condensation and we can ask for the effective glueball superpotential. In [2] it was proposed that the effective superpotential is computed by a unitary matrix model. If one includes hypermultiplets, the effective superpotential should be computed by a matrix quantum mechanics.

As a test of these ideas, one may try to recover the Seiberg-Witten curve for the (compactified) five-dimensional gauge theory by turning off the superpotential. The curve has already been obtained through many other approaches, such as integrable systems [3], (p,q) five-branes [4,5], geometric engineering [6,7], deconstruction [8] and instanton calculus [9], and the links between many of these calculations are known. In section 2 we will review how the curve is obtained through geometric engineering.

The routes to and through the relevant matrix model can be made to follow in rough outline the routes that have been taken for the four-dimensional case. We choose to focus on some of the earlier ideas, which emphasize the embedding of the gauge theory in a larger string theory. In this spirit, in section 3 we consider a D-brane setup for the five dimensional Yang-Mills theory where the computation of F-terms can be argued to reduce

to a matrix model. The effective superpotential that the matrix model computes arises from an open-closed duality in string theory where the D-branes dissolve into flux. In section 4 we analyse the matrix model and try to recover the Seiberg-Witten curve in the limit where the superpotential is switched off. The results in the latter part of section four are somewhat puzzling. Besides the expected critical point of the glueball superpotential we seem to find some additional critical points. These can presumably be eliminated by a more careful analysis.

One of the main differences with the four-dimensional case arises from the possibility of a Chern-Simons term in five dimensions. The stringy picture gives us a clear intuition for how to deal with this. Basically there is an integer ambiguity in the glueball superpotential which is related to the Chern-Simons level. It should be possible to derive this directly from the matrix model, but we will not attempt this here. We will also suggest a generalisation of matrix model methods to theories with more general non-minimal kinetic terms for the gauge fields.

A matrix model for a five dimensional gauge theory with a non-hyperelliptic curve has been discussed in [10]. Recently extensions to six dimensional gauge theories have been discussed in [11] and [12].

2. A brief tour of geometric engineering

This section is based on [13,6,14,15,16].

2.1. Four-dimensional geometric engineering

Geometric engineering starts with the observation that IIa string theory on an ALE space gives rise to a six dimensional gauge theory with sixteen SUSY localised near the 'origin' of the ALE. The equation for an ALE is

$$uv + x^n = 0. \tag{2.1}$$

The singularity may be resolved and the middle dimensional homology is then isomorphic to the root lattice of the algebra of $SU(n)$; it is spanned by $n - 1$ 2-spheres with volume $|a \cdot \alpha|$ where a is the complexified Kähler form on the ALE, and α runs over the simple roots of the $SU(n)$ algebra. The non-abelian gauge bosons are identified with D2 branes wrapped on various two-spheres in this geometry and their mass is proportional to the volume of the 2-sphere they are wrapped on.

In order to get a four dimensional gauge theory with eight supercharges and no adjoint matter, one fibers the ALE geometry over a 2-sphere in such a way that the total three-fold is a Calabi-Yau space. The four dimensional gauge coupling (squared) is then inversely proportional to the area A of the base S^2:

$$\frac{1}{g_4^2} \sim \frac{A}{g_s \ell_s^2}. \tag{2.2}$$

Further, a massless four-dimensional vector multiplet contains an adjoint-valued complex scalar Φ. It may be expanded in terms of the $SU(n)$ Cartan generators T_i as $\Phi = a_i T^i$, in other words, the a_i parametrise the Coulomb branch.

String worldsheets wrapped on the 2-spheres act as (point-like) gauge instantons; this may be seen either from type II on K3/heterotic on T^4 duality, or from the fact that in type II compactified on K3 we have the six-dimensional equation $d * e^{-2\phi} H = -\text{Tr} F \wedge F$. The instanton number is the number of times the worldsheet configuration is wrapped on the base $\mathbf{P}^1$.

The low energy effective action (up to two derivatives) for the massless degrees of freedom of string theory on this background is determined by a prepotential $\mathcal{F}(a_i, A)$ and may be computed by an application of mirror symmetry. After taking a decoupling limit which keeps an interacting gauge theory but throws away all other modes, one recovers the gauge theory prepotential and action:

$$4\pi S = \int d^4\theta \left(\frac{\partial \mathcal{F}}{\partial a_i} \bar{a}_i - \frac{\overline{\partial \mathcal{F}}}{\partial a_i} a_i \right) + \frac{1}{2} \int d^2\theta \, \frac{\partial^2 \mathcal{F}}{\partial a_i \partial a_j} W_{i,\alpha} W_j^\alpha + h.c. \tag{2.3}$$

From the gauge theory perspective, the prepotential consists of three pieces: $\mathcal{F} = \mathcal{F}_{\text{tree}} + \mathcal{F}_{1-\text{loop}} + \mathcal{F}_{\text{instanton}}$, where

$$\mathcal{F}_{\text{tree}} = \frac{1}{2} \tau_0 \, a^t \cdot C \cdot a$$

$$\mathcal{F}_{1-\text{loop}} = \frac{i}{8\pi} \sum_{\text{roots } \alpha} (\alpha \cdot a)^2 \log \frac{(\alpha \cdot a)^2}{\Lambda^2} - \frac{i}{8\pi} \sum_{\text{flavours}} \sum_{\text{weights } \mathbf{w}} (\mathbf{w} \cdot a)^2 \log \frac{(\mathbf{w} \cdot a)^2}{\Lambda^2} \tag{2.4}$$

$$\mathcal{F}_{\text{instanton}} = \sum_{k>0} \Lambda^{Tk} \mathcal{F}_k(a_i).$$

We used $\tau_0 = \frac{4\pi i}{g_0^2} + \frac{\theta}{2\pi}$ for the bare gauge coupling, C for the Cartan matrix, and T for the instanton weight which depends on the gauge group and the matter representations (it's $2N_c - N_f$ for $SU(N_c)$ with N_f fundamental matter). From the stringy point of view these

three pieces receive contributions from classical string theory (here Calabi-Yau intersection numbers), genus zero worldsheet instantons with zero gauge instanton number with a small piece from the classical intersection numbers, and genus zero worldsheet instantons with gauge instanton number k respectively.

For many gauge theories the relevant Calabi-Yau geometry is actually toric, and the whole procedure may be automated. Let us discuss the case of pure $SU(n)$ gauge theory. The linear sigma model has gauge group $U(1)^{n-1}$ and $n+2$ chiral fields X_j. The charge vectors are given by

$$\begin{aligned}
q_b &= (1, 1, -2, 0, 0, 0, 0, \ldots, 0, 0, 0, 0) \\
q_{f_1} &= (0, 0, 1, -2, 1, 0, 0, \ldots, 0, 0, 0, 0) \\
q_{f_2} &= (0, 0, 0, 1, -2, 1, 0, \ldots, 0, 0, 0, 0) \\
\vdots &= \quad \vdots \\
q_{f_{n-2}} &= (0, 0, 0, 0, 0, 0, 0, \ldots, 1, -2, 1, 0) \\
q_{f_{n-1}} &= (0, 0, 0, 0, 0, 0, 0, \ldots, 0, 1, -2, 1)
\end{aligned}$$

(2.5)

The sum of the charges in each row is zero; this is the Calabi-Yau condition. The columns of this matrix can be identified with divisors D_j defined by $X_j = 0$. For each row there is a D-term, whose solution set contains a $T^* \mathbf{P}^1$ with a minimal size $\mathbf{P}^1$ with a volume determined by the corresponding FI-term. Thus to each row we can associate a curve C_i, namely the minimal size $\mathbf{P}^1$. The first vector corresponds to the base, and the remaining ones correspond to the exceptional curves in the ALE. The (i, j)th entry of the matrix is the intersection number $C_i \cdot D_j$. Note that the last $n-1$ vectors contain the negative of the Cartan matrix for A_{n-1}, since it describes the intersections of $\mathbf{P}^1$'s in the ALE.

For the purpose of getting a graphical representation of the data and finding the mirror, one introduces a maximal set of integer vectors ν_i such that

$$\sum q \cdot \nu = 0. \tag{2.6}$$

In the present case we can take

$$\nu = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & n-1 \end{pmatrix} \tag{2.7}$$

4

The Calabi-Yau condition implies we can always take the first column to be all 1's. The rows ν_i are usually represented graphically by forgetting the first entry and drawing a node in a plane with position as indicated by the remaining entries. The resulting toric diagram is sometimes also called a 'grid diagram.'
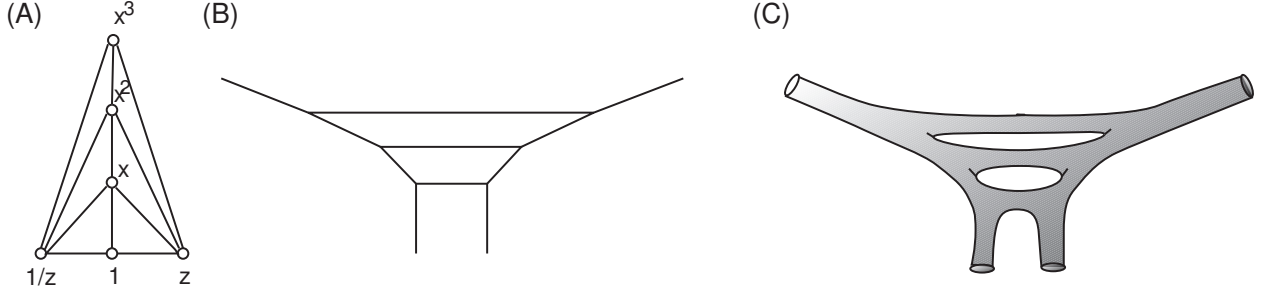

fig. 1: (A): Toric diagram from the data (2.7) with $n = 3$. This corresponds to pure $SU(3)$ gauge theory. (B): Dual web diagram. (C): The Seiberg-Witten curve from thickening the web diagram.

The mirror geometry is determined by associating a monomial $z^i x^j$ to a node with coordinates (i, j) in the toric diagram. The resulting equation is

$$b_1 z + \frac{b_{-1}}{z} + \sum_{i=0}^{n} c_i x^i = 0. \tag{2.8}$$

One may think of this curve as the 'reduction' of the non-compact Calabi-Yau three-fold

$$b_1 z + \frac{b_{-1}}{z} + \sum_{i=0}^{n} c_i x^i = u\, w. \tag{2.9}$$

Reduction means that the periods of the holomorphic (3,0) form $\Omega = dz dx du/(zxu)$ over 3-cycles in the Calabi-Yau can be written as periods of the (1,0) form $\log(x) dz/z$ over 1-cycles of (2.8).

The independent complex structure parameters in (2.8) can be read from the matrix q (2.5):

$$z_{f_i} = \frac{c_{i-1} c_{i+1}}{c_i^2}$$
$$z_b = \frac{b_{-1} b_1}{c_0^2}. \tag{2.10}$$

If we denote the (complexified) volumes of each of the $\mathbf{P}^1$'s associated to the charge vectors as t_k, then we have $t_k = -\log(z_k)$. It is convenient to introduce flat Kähler coordinates T_k which measure physical quantities such as the exact masses of W-bosons. On the mirror

side, the exact masses of BPS particles are computed by period integrals of $\log(x)dz/z$. These periods satisfy Picard-Fuchs differential equations. One can then find the T_k explicitly by solving differential equations and matching the log terms:

$$T_k = -\log(z_k) + \sum_{n \geq 0} g_n z_b^{n_b} z_{f_i}^{n_i}. \tag{2.11}$$

The coordinates T_k and t_k differ by instanton corrections:

$$T_k = t_k + \sum k_n e^{-n_b t_b} e^{-n_i t_{f_i}}. \tag{2.12}$$

Finally we need to take a limit $\alpha' \to 0$ in which the gauge theory decouples from other modes but remains interacting:

$$\exp(-1/g_4^2) \sim \exp(-T_b) \sim \epsilon^{2n} \Lambda^{2n}, \quad T_{f_i} \sim \epsilon(e_i - e_{i+1}) \quad \text{as } \epsilon \sim g_s 2\pi \sqrt{\alpha'} \to 0. \tag{2.13}$$

In this limit m_W/Λ remains finite, $e_i = a \cdot \lambda_i$ where λ_i are the fundamental weights of $SU(n)$, and we take a_i to be the quantum corrected Coulomb branch parameters (so that $e_i - e_j$ measure exact masses of W-bosons). Then we may write the above equation in terms of physical parameters as

$$z + \frac{\epsilon^{2n} \Lambda^{2n}}{z} + \prod_{i=1}^{n} (x - \exp(\epsilon e_i)) = 0. \tag{2.14}$$

Putting $x = e^{\epsilon v}$ and taking the limit $\epsilon \to 0$, we obtain

$$z + \frac{\Lambda^{2n}}{z} + \prod_i (v - e_i) = z + \frac{\Lambda^{2n}}{z} + P_n(v) = 0. \tag{2.15}$$

Defining $y = 2z + P_n(v)$, this may be written in hyperelliptic form:

$$y^2 = P_n(v)^2 - 4\Lambda^{2n}. \tag{2.16}$$

Moreover one recovers the expected Seiberg-Witten differential:

$$\lambda = \log(x) \frac{dz}{z} \to R v \frac{dz}{z}. \tag{2.17}$$

This differential can be used to recover the exact prepotential. Namely if we choose a basis of A-cycles for the curve such that $\int_{A_i} v \frac{dz}{z} = a_i$ and we define the dual B-cycles by $A_i \cdot B_j = \delta_{ij}$, then we have $\frac{\partial \mathcal{F}}{\partial a_j} = \int_{B_j} v \frac{dz}{z}$.

6

One can introduce fundamental matter by performing a toric blow-up of a point on the base $\mathbf{P}^1$. This adds a new $\mathbf{P}^1$ for which the homology class acts as a weight of the fundamental representation of $SU(n)$, say the highest weight λ_+. The other fundamental weights can be obtained by adding classes corresponding to the $SU(n)$ roots. The resulting toric diagram admits several triangulations. We have drawn two of them below:
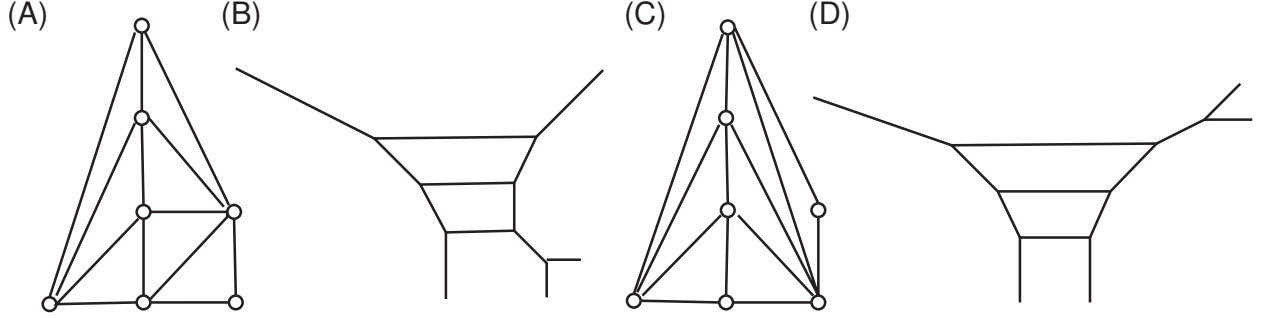

fig. 2: (A): Toric diagram for $SU(3)$ with a fundamental hypermultiplet. (B): Dual web diagram. (C): A different triangulation of the diagram in (A). (D): Dual web diagram for (C).

There is an extra parameter corresponding to the mass of the hypermultiplet m. If we take a D2-brane wrapped on the $\mathbf{P}^1$ with class λ_+ and volume t_m, we will get a particle in four dimensions with mass $|\lambda_+ \cdot a + m|$, so we set

$$t_m = -\log(z_m), \quad T_m = R(\lambda_+ \cdot a + m). \tag{2.18}$$

This will yield the Seiberg-Witten curve

$$z + \frac{\Lambda^{2n-1}(v+m)}{z} + \prod_i (v - e_i) = 0. \tag{2.19}$$

To any given toric diagram one can associate a dual diagram by drawing the bi-sectors for a given triangulation as in fig. 1B, fig. 2B and D. This is often called a (p,q) web diagram. One can think of the toric Calabi-Yau space as the total space a non-trivial T^2-fibration over $\mathbf{R}^2$ together with a trivial $\mathbf{C}$-fibration. If we identify the $\mathbf{R}^2$ with the plane of the dual diagram, then the web tells us where the T^2-fibration degenerates. Namely at each vertex the full T^2 shrinks to a point, and at an edge with tangent direction (p,q) (where p and q are integers and the vector is primitive) the one-cycle with class (p,q) in T^2 shrinks to a point. We can use dualities to map this to the corresponding web of (p,q) five-branes in IIb string theory, where (p,q) stands for the (R,NS) charge of the five-brane associated to the edge. The web picture is often a useful way to think about the Calabi-Yau.

Strictly the (p,q) fivebranes give rise to five dimensional gauge theories, but we can get four dimensional theories by compactifying the worldvolumes of the branes on a circle and shrinking it to zero size. For finite radius of the circle we can picture this by thickening the legs of the web diagram, as in fig. 1C. The resulting curve is actually the Seiberg-Witten curve.

We may think of the horizontal lines in fig. 1B as D5-branes. When they coincide we get an enhanced gauge symmetry; we denote their vertical coordinate by $e_i \ell_s^2$ (which we can make into a complex number by adding the Wilson line around the S^1 that shrinks to zero size). Then the W-bosons correspond to strings stretching between the D5-branes and their masses are proportional to $|e_i - e_j|$.

Matter multiplets are introduced by adding semi-infinite D5-branes, as in fig. 2B and D, and quantising strings stretched between the gauge and the matter branes. The vertical coordinate of the matter brane is related to the real part of the mass of the hypermultiplet (again the imaginary part comes from a Wilson line).

2.2. Five-dimensional geometric engineering

The appearance of $\epsilon = g_s \ell_s = R_{11}$ above is very suggestive. Indeed it turns out that it is more natural to consider geometric engineering of five dimensional gauge theories through M-theory on a Calabi-Yau, and recover the four dimensional theories as a limit $R_{11} \to 0$ of these gauge theories. This is also hinted at by the (p,q) five-brane description.

The reasoning is similar. M-theory on an ALE space again gives rise to non-abelian gauge symmetry (the non-abelian gauge bosons are now wrapped M2 branes) and fibering over a $\mathbf{P}^1$ as before yields gauge theories in five dimensions with eight supercharges. The action in five dimensions is still determined by a prepotential which is a function of the vector multiplet moduli ϕ^i (these are now real scalars). The major distinction with the four dimensional case is that the prepotential has a classical part and a one-loop contribution, but does not receive instanton corrections. The triple intersection numbers of the Calabi-Yau, when expressed in terms of physical quantities, reproduce exactly the expected prepotential [14]:

$$\mathcal{F} \sim \frac{1}{2g_0^2} h_{ij} \phi^i \phi^j + \frac{k_{\rm cl}}{6} d_{ijk} \phi^i \phi^j \phi^k + \frac{1}{12} \sum_{\text{roots } \alpha} |\alpha \cdot \phi|^3 - \frac{1}{12} \sum_f \sum_{\text{weights } \mathbf{w}_f} |\mathbf{w} \cdot \phi + m_f|^3. \quad (2.20)$$

Here $h_{ij} = \text{Tr}(T_i T_j)$ and $d_{ijk} = \frac{1}{2}\text{Tr}(T_i(T_j T_k + T_k T_j))$ with T_i the Cartan generators. Moreover k corresponds to the Chern-Simons level in five dimensions:

$$\frac{k_{\text{cl}} d_{ijk}}{24\pi^2} \int_{R^5} A^i \wedge F^j \wedge F^k. \tag{2.21}$$

When we compactify the five-dimensional theory on a circle, the prepotential receives additional corrections. We can interpret these corrections as one-loop corrections due to a Kaluza-Klein tower of BPS particles, or alternatively one can interpret the corrections as a one-loop term in five dimensions plus worldline instantons of the five dimensional particle. For each particle in five dimensions with mass m we get a contribution

$$\tau = \frac{\partial^2 \mathcal{F}}{\partial \phi^2} \sim \frac{1}{2\pi i} \log\left(\frac{1}{R^2 \Lambda_0^2} \sinh^2(Rm/2)\right) \tag{2.22}$$

with $1/g_4^2 = 2\pi R/g_5^2$. Note that if we put $R\Lambda_0 \sim e^{RM_0/2}$, we recover the usual five dimensional linear divergence for $1/g_5^2$ as $R \to \infty$.

Since the dilaton lives in a neutral hypermultiplet and does not talk to the vector multiplets, we can use mirror symmetry to compute the full prepotential. The crucial difference with the four dimensional case is that now when we take the decoupling limit $l_p \to 0$, we do not have to take a special scaling limit of the curve. Instead we have

$$\exp(-T_b) \sim R^{2n}\Lambda^{2n}, \quad T_{f_i} \sim RM_{W,i}. \tag{2.23}$$

The resulting Seiberg-Witten curve is of the form

$$z + \frac{\Lambda^{2n}}{z} + \det(x - U) = 0 \tag{2.24}$$

with Seiberg-Witten differential

$$\lambda_{SW} = \log(x)\frac{dz}{z}. \tag{2.25}$$

The relation between the coefficients in this equation and the physical parameters can be determined by solving Picard-Fuchs equations, with the leading terms at large volume fixed by (2.11).

Actually (2.24) is only one of the curves one could get. The gauge and matter content fixes the toric Calabi-Yau up to a small ambiguity in the triple intersection numbers [14]. This ambiguity corresponds to k_{cl} in (2.20) and gives rise to the Chern-Simons term. All these geometries give rise to the same Seiberg-Witten curve when we take $R \to 0$. The

9

toric diagram we discussed before (fig. 1A) corresponds to level n, as one can check for instance by computing the triple intersection numbers of the Calabi-Yau. The other toric diagrams for the case of $SU(3)$ gauge theory are given below.
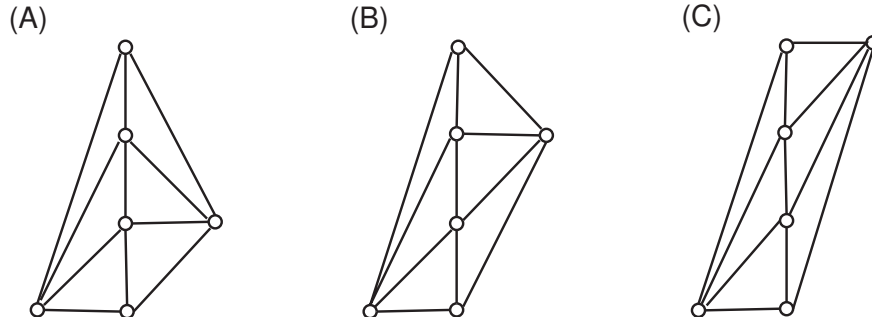

fig. 3: Toric diagrams for $SU(3)$ gauge theory with (A): level 2, (B): level 1, and (C): level 0.

There is a simple way to derive the diagrams for different levels from any given one by adding hypermultiplets. The reason is that integrating out a hypermultiplet with a large mass m is well known to give a contribution $\frac{1}{2}\frac{\text{Re}(m)}{|\text{Re}(m)|}$ to the Chern-Simons term [17]. Varying the mass between plus and minus infinity interpolates between successive levels. An example is given in fig. 2: as we vary the vertical position (i.e. the real part of the mass) of the semi-infinite D5-brane, we cross all of the compact D5-branes which results in 'flops' (changes in the triangulation of the toric diagram). When we take the semi-infinite D5-brane off to infinity we can eliminate a node in the toric diagram. This interpolates between level 3 and level 2 for $SU(3)$.

For general level k and one hypermultiplet, the curve may be written as

$$z + \frac{(R\Lambda)^{2n-1}e^{Rm/2}x^{n-k_c-1/2}(x - e^{-Rm})}{z} = P_n(x) \qquad (2.26)$$

with $P_n(x) = \det(x - U)$ and Seiberg-Witten differential $\lambda_{SW} = \log(x)d\log(z)$. Here k_c is the classical CS level of the five dimensional gauge theory; there is also a one-loop quantum contribution $\frac{1}{2}\frac{\text{Re}(m)}{|\text{Re}(m)|}$ from the matter field. For the case at hand we will put $k_q = k_c + \frac{1}{2}$ which must be an integer to avoid anomalies. A space-time parity transformation corresponds to

$$k_c \to -k_c, \ m \to -m, \ U \to U^{-1}. \qquad (2.27)$$

This transformation gives us back the same curve, expressed in the variable $\tilde{x} = 1/x$ and with a redefined z.

Finally for $R|m|$ large compared to $\log(R|\Lambda|)$ we can put $(R\tilde{\Lambda})^{2n} = (R\Lambda)^{2n-1}e^{\pm Rm/2}$ and take the limit $m \to \pm\infty$ in (2.26). Varying the mass of the hypermultiplet allows us to interpolate between different CS levels in (2.26).

10

3. Brane/flux duality

An alternative realisation of five dimensional Yang-Mills is through D-branes. For the case of pure gauge theory with eight supercharges we can consider type IIa string theory on an A_1 ALE space times $\mathbf{C}^*$ and wrap D6-branes on the $S^2 \times S^1$.[1] The open string description leads to a matrix model, and the planar diagrams of this matrix model compute an effective superpotential. The same effective superpotential also arises from the dual closed string description. We use this setup to discuss some aspects of the matrix model and effective superpotential.

For a single D6, the moduli space is parametrised by the position on the cylinder (the real scalar in the vector multiplet) and a Wilson line (parametrised by the dual of the S^1). In the following it will frequently be useful to apply a T-duality on the S^1. In particular the modulus of a D6-brane is described by the position of a D5-brane on the T-dual type IIb cylinder. For N D6-branes, the moduli space is parametrised by the position of N D5-branes on the IIb cylinder, the eigenvalues of the complexified Wilson line[2]

$$U = P \exp(\int dx_5 \, A_5 + \varphi). \tag{3.1}$$

The Chern-Simons level arises in the following way. The D-brane action depends on background flux $F_{RR}^{(2)}$ through the WZ term $\int e^F \wedge \sum C_i$. For a each D6 wrapped on $\mathbf{P}^1$, we have

$$\int_{D6} F \wedge F \wedge F \wedge C^{(1)} = -\int_{R^{3,1} \times S^1} A \wedge F \wedge F \int_{S^2} dC^{(1)} = -k \int_{R^{3,1} \times S^1} A \wedge F \wedge F. \tag{3.2}$$

One may introduce fundamental matter by adding extra D6 branes wrapped on certain non-compact cycles. If the superpotential is turned off then the matter D6 brane wraps the S^1 on $\mathbf{C}^*$ and a cycle of the form $x_1 = 0$ in the ALE

$$|x_1|^2 + |x_2|^2 - 2|x_3|^2 = r \ /U(1). \tag{3.3}$$

This cycle intersects the $\mathbf{P}^1$ given by $x_3 = 0$ in precisely one point inside the ALE. The separation on $\mathbf{C}^*$ with the centre-of-mass of the stack of N D6 branes is related to the

---

[1] Of course this is related to the set-up of the previous section through dualities. A D6 wrapped on $\mathbf{P}^1$ inside an ALE is essentially T-dual to a D5-brane stretched between to NS5-branes, which gives the (p,q) five-brane picture. On the other hand as we mentioned in the previous section, this is dual to geometric engineering.

[2] For convenience we set $R = 1$ in this section

(real part of the) mass of the hypermultiplet, and the difference between the Wilson lines on the S^1 is related to the imaginary part of the mass.[3] Note that such matter D6 branes act as domain walls for the background flux $F_{RR}^{(2)}$ on $\mathbf{P}^1$. Therefore they affect the CS level depending on the hypermultiplet mass, and one can change the value of the background flux by moving matter D6-branes between $\pm\infty$, hence change the CS level of the field theory.

Superpotential terms for the pure gauge theory theory living on the D6 branes may be computed through $U(N)$ Chern-Simons theory on $S^2 \times S^1$:

$$S_{CS} = \frac{i\kappa}{4\pi} \int_{S^2 \times S^1} \mathrm{Tr}\, A \wedge dA + \frac{2}{3} A \wedge A \wedge A. \qquad (3.4)$$

We write the gauge fields as $A_z, A_{\bar{z}}$ and A_t, where t is the coordinate along the S^1 and $z, \bar{z}$ parametrise the S^2. Integrating out gauge degrees of freedom has the effect of shifting $\kappa \to \kappa + N$. We will require A_t to be independent of t, and identify $g_s = 2\pi i/(\kappa + N)$. However if there are matter D6 branes or k units of flux through the $\mathbf{P}^1$ then the action needs some modification. Let us first consider a single matter brane with zero background flux. Strings stretching between the matter and gauge D6-branes were considered in the topological set-up in [18] and they give rise to a complex scalar ϕ. Denoting the coordinate along the S^1 by t, the action is modified to

$$S = S_{CS} + \int_{S^2 \times S^1} \bar{\phi}(\partial_t + A_t + m)\phi\, \delta^{(2)}_{x_1 = x_3 = 0}. \qquad (3.5)$$

The exact normalisation of this term is not relevant here, since we can rescale $\phi, \bar{\phi}$.

Since we can interpolate between different Chern-Simons levels by varying the mass of the matter multiplet, let us see how this modifies the action. Integrating out ϕ gives the following term in the action:

$$-\log\det\left(\partial_t + A_t + m\right) = -\mathrm{Tr}\log\sinh(\pi(A_t + m)) + \text{constant}. \qquad (3.6)$$

As $m \to \infty$ we kill the $e^{-x/2}$ term in sinh, and as $m \to -\infty$ we kill the $e^{x/2}$ term. Therefore by starting with pure Chern-Simons theory and moving branes from minus to plus infinity, we see that we can account for the background flux by adding a term

$$-k \int_{S^2 \times S^1} \mathrm{Tr}\, A_t\, \omega \qquad (3.7)$$

---

[3] There are corrections e.g. due to the warp factor that arises in the presence of space-filling branes. It is better to think of it as an unobservable bare mass.

to the action, with ω a (1,1) form with unit volume on $\mathbf{P}^1$.[4] The periodicity of A_t implies that k is quantised.

We can also turn on a superpotential $\mathrm{Tr}\, W(U)$ in the gauge theory. One can incorporate this by turning on the corresponding term in the Chern-Simons theory, with $U = P \exp \int dt\, A_t$. The geometric effect of the superpotential is most easily understood in the T-dual type IIb picture, where it deforms the trivial fibration of the ALE over the cylinder to a non-trivial one according to the equation

$$uw = (y - W'(x))(y + W'(x)). \tag{3.8}$$

Here $x = e^{Rv}$ and the prime denotes differentiation with respect to v. Therefore altogether we consider the partition function for the theory with action

$$S = S_{CS} + \int_{S^2 \times S^1} dt\, \bar{\phi}(\partial_t + A_t + m)\phi\, \delta^{(2)}_{x_1 = x_3 = 0} - k\, \mathrm{Tr}\, A_t\, \omega - \frac{1}{g_s} \mathrm{Tr}\, W(U)\, \omega \tag{3.9}$$

Note that $A_z, A_{\bar{z}}$ appear only in S_{CS} and appear linearly. Integrating out $A_{\bar{z}}$ gives $F_{tz} = 0$ and reduces S_{CS} to $\int \mathrm{Tr} A_z \partial_{\bar{z}} A_t$. After gauge fixing, this looks like the topological part of the D5-brane action considered in [1], and is consistent with the form of the geometry (3.8) and the factor of $1/g_s$ in front of the superpotential in (3.9). Integrating out A_z sets $\partial_{\bar{z}} A_t = 0$ which implies that A_t is a constant (we already got rid of the t-dependence). Therefore we obtain a matrix quantum mechanics, and the measure for A_t is that of a unitary matrix model due to the periodic identifications. The partition function is

$$Z = \frac{1}{\mathrm{Vol}(U(N))} \int dU d\phi(t)\, e^{-\frac{1}{g_s} \mathrm{Tr}\, W(U) - k\mathrm{Tr}\, \log U + \int_{S^1} dt\, \bar{\phi}(\partial_t + A_t + m)\phi} \tag{3.10}$$

The geometry (3.8) contains finite size $\mathbf{P}^1$'s obtained from blowing up the singular points, which are in one-to-one correspondence with the critical points of W. The D5-branes we get from T-dualising the D6 branes wrap these $\mathbf{P}^1$'s. In the gauge theory the superpotential Higgses the $U(N)$ to $U(N_1) \times \ldots U(N_n)$ where $\sum N_i = N$ and N_i is the number of eigenvalues of U that we put at the ith critical point of W. In the present set-up this corresponds to wrapping N_i D5-branes on the $\mathbf{P}^1$ we get from blowing up the ith singular point of (3.8).

---

[4]   The equation of motion for A_t then imply that the gauge field has k units of flux through the S^2. This is similar to the baryon vertex for k quarks where the quarks are strings stretching from the gauge D6 branes to infinity.

We can alternatively describe this system as a closed string theory on a Calabi-Yau which is obtained by excising each $\mathbf{P}^1$ and glueing in an S^3. In the dual geometry the D-branes are replaced by fluxes. From the perspective of the field theory this is a large N dual description in terms of glueball fields $S_i \sim \mathrm{Tr}_{\mathrm{U(N_i)}} W_\alpha W^\alpha$. The scalar component of the superfield S_i is related to the size of the ith S^3 in the dual geometry and therefore corresponds to a complex structure parameter of the dual geometry.

In the closed string set-up the Calabi-Yau three-fold is given by[5]

$$uw - y^2 + W'(x)^2 + q(x) = 0 \tag{3.11}$$

where $q(x)$ corresponds to a (log-)normalisable complex structure deformation of (3.8) which gives finite volume to the S^3's. The complex structure parameters are governed by an effective superpotential[6]:

$$W_{\mathrm{eff}} \sim \int_{\mathrm{CY}_3} H \wedge \Omega \tag{3.12}$$

where $H = H_{RR} + \tau H_{NS}$ and Ω is the holomorphic 3-form on the Calabi-Yau on the IIb side. One can effectively reduce the superpotential to[7]

$$W_{\mathrm{eff}} \sim \int_\Sigma h \wedge \lambda \tag{3.13}$$

where $\lambda = \sqrt{W'(x)^2 + q(x)}\, dx/x$, and Σ is the curve defined by

$$y^2 = W'(x)^2 + q(x). \tag{3.14}$$

A representation of this curve is given fig. 4.

---

[5] More accurately, the space-time metric also includes a warp factor that takes into account the backreaction of the non-zero fluxes.

[6] Physically to find the effective superpotential we should expand around the space-time solution which includes the warp factor and integrate out all unwanted fields. The standard claim in the literature (eg. [19]) is that the effective superpotential is defined on the moduli space of the original Calabi-Yau before any fluxes were turned on. This requires a remarkable cancellation of the warp factor. I would like to thank V. Zhukov for a conversation on this long ago.

[7] The correct normalisation should be $W_{\mathrm{eff}} = -\frac{1}{2} \int_\Sigma h \wedge \lambda$ so that in the $g_s N \to 0$ limit (in which $q(x) \to 0$) we get the expected superpotential $\sum N_i W(x_i) + C$, with x_i the critical points of W and C depends on the cut-offs $\Lambda_\pm$ but not on the x_i. Nevertheless we will not explicitly include the $-\frac{1}{2}$ below.

Expanding W_{eff} in terms of the periods of h and λ gives us an effective superpotential for the glueball fields S_i. It is not entirely clear we can interpret it as an effective superpotential in the conventional sense since we are in a massive vacuum and the S_i may not be the lightest degrees of freedom. However it clearly has non-trivial information because we use it to compute exact tensions of BPS domain walls. Moreover after minimising, solving for the VEV's of the S_i and substituting back, we can interpret the resulting superpotential as a generating function for the quantum expectation values $\langle \text{Tr}(U^m) \rangle$.[8] The reduction of the open string theory to the matrix quantum mechanics (3.10) implies that the effective superpotential can be computed from the matrix quantum mechanics.

In order to expand (3.13), we would like to know the periods of h. This works similar to the four dimensional case, with a twist because the Riemann surface has two extra punctures. The integral of h around each cut should give the number of eigenvalues we put at the corresponding critical point of W. What about the other periods?
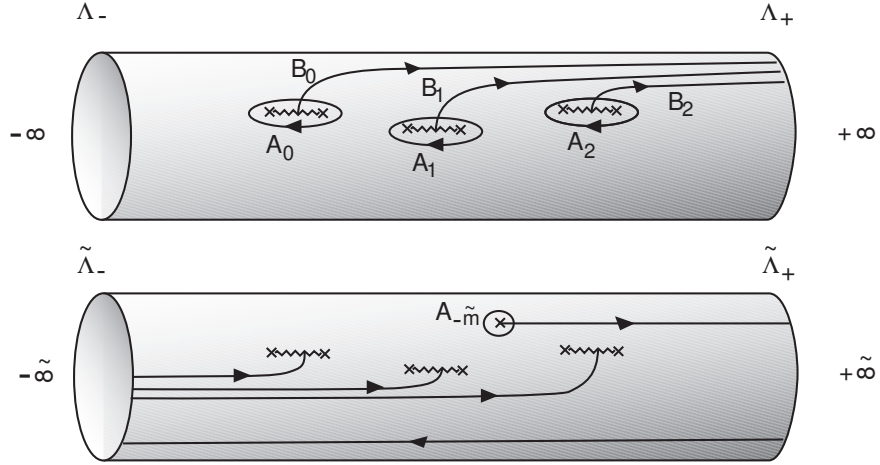

fig. 4: The matrix model curve $y^2 = W'(x)^2 + q(x)$ as the double cover of a cylinder parametrised by $x = e^{Rv}$. The B-cycles essentially follow flux lines which are discussed in the text. The geometry of the curve looks very similar to a Seiberg-Witten curve as in fig. 1C turned 90^0, but they are unrelated at this point.

After T-duality, the background flux $F^{(2)} \sim k\,\omega_{S^2}^{(2)}$ gives rise to a background 3-form flux. Upon reduction to Σ, this gives rise to a background flux for h. Similarly any D6 brane is a magnetic source for 2-form flux; after T-duality the matter D6 branes gives

---

[8]   Actually the effective superpotential refers to a particular UV completion of the field theory. See [20,21] for a discussion.

rise to additional sources for 3-form flux, which in turn gives pointlike sources for h after reduction. Therefore the flux h has a constant term k related to the CS level and sources at the position of the matter D5-branes.

The background flux and the matter sources appear on the 'x-plane' (cylinder); the algebraic curve obtained from reduction of the Calabi-Yau is a double cover of that. So we should decide on which sheet to put the matter sources and the background flux for h. Let us first understand how it works for zero background flux and no matter.

For N gauge D5-branes we get N charged particles (or matrix model eigenvalues) on the cylinder. The fluxes through $\pm\infty$ on each sheet must therefore add up to $\pm N$. On the upper sheet we have to make a choice to distribute f units of flux at $-\infty$ and $N - f$ units of flux at $+\infty$. No physical quantities will depend on this choice, but once we make it all other fluxes will be fixed. We will choose $f = 0$. Then the gauge D5-branes are effectively attracted to a particle at $+\infty$ with charge N. In order to balance this force, in the absence of background flux or other sources, on the lower sheet we must then put the $-N$ units of flux entirely at $-\tilde{\infty}$. Not doing this would lead to an additional force term in the matrix model action.

Now we introduce matter branes. Recall that classically we have $\langle \text{Tr}(U^r) \rangle = \oint x^r h(x) dx$ where h has poles at positions corresponding to the classical positions of the N gauge D5 branes, and is independent of the position of additional matter branes (if the hypers don't get a VEV). Therefore h should get contributions from matter only on the 'quantum' sheet that is not visible classically, and we take this to be the lower sheet.[9] Then the background flux k will also only appear on this sheet, because we can create it by moving matter branes around on the second sheet. This fixes all the A-periods of h.

Finally let us discuss the B-periods of h. As in the four-dimensional case we will take the compact B-periods to be zero (although this can be generalised slightly as in [22]). The non-compact periods are divergent however so we need to introduce cut-offs to define the integrals. In principle there are four asymptotic infinities and we need to introduce a cut-off for each of them. This might seem very strange from the gauge theory point of view where we need only one cut-off. However it is by now a familiar story that the gauge theories we are analysing are not renormalisable and the matrix model seems to refer to a particular UV completion [20,21]. Recently it was understood in the four dimensional case [21] that introducing cut-offs for the divergent integrals in the effective superpotential

---

[9]   If some of the hypers get a VEV, the poles do appear on the upper sheet [22].

is indistinguishable from introducing sufficiently many matter branes on each sheet at the position of the cut-offs, so that the total matter content is $\tilde{N}_f = 2N_c$. In other words, one can think of the cut-offs as the masses of additional matter multiplets. It appears that this explanation is valid in the five-dimensional case also. Indeed we have already observed that we effectively have N_c charges at infinity on the upper sheet and $N_c - N_f$ charges at infinity on the lower sheet (for any Chern-Simons level). We can identify these charges with additional matter branes, thus embedding the original theory into a theory with total number of flavours $\tilde{N}_f = 2N_c$. Interestingly, the underlying $\mathcal{N} = 2$ theories are also precisely the theories that can be lifted to six dimensions, which suggests that should be the natural starting point.

At any rate, we will use a cut-off $\Lambda_+, \tilde{\Lambda}_+$ for $+\infty, +\tilde{\infty}$ and $\Lambda_-, \tilde{\Lambda}_-$ for $-\infty, -\tilde{\infty}$. Due to the special form of the A-periods of λ that we will derive later, it turns out that the glueball superpotential depends only on the combination

$$\int_{-\tilde{\infty}}^{-\infty} h + \int_{+\tilde{\infty}}^{+\infty} h. \tag{3.15}$$

To get the correct running of the gauge coupling it appears that we need

$$\frac{1}{2}\int_{-\tilde{\infty}}^{-\infty} h + \frac{1}{2}\int_{+\tilde{\infty}}^{+\infty} h = \tau \tag{3.16}$$

which is consistent with the $R \to 0$ limit.

4. The matrix model.

In the last section we have indicated from the string set-up that we expect to be able to compute F-terms of the five-dimensional gauge theory in terms of a 'dimensionally reduced' theory. One can also argue directly from the field theory perspective that such a reduction occurs [2].

We are interested in the F-terms, which according to [23] can be written is 4D super-space as

$$\int d^5x \int d^2\theta \, \mathrm{Tr} \, (\tau + k\Phi)W_\alpha W^\alpha + H^c(\partial_t + \Phi + m)H + \mathrm{Tr} \, W(U) + \mathrm{h.c.} \tag{4.1}$$

Now the logic of [2] runs as follows: if we think of the coordinate t as a label, then we can think of the 5D theory as a 4D theory with an infinite number of fields. Then applying

what we know about computing F-terms in a 4D theory, we expect that we need to consider a bosonic quantum mechanics with action given by

$$\mathcal{S} = -\frac{1}{Rg_s} \int dt \, [H^c(\partial_t + \Phi + m)H + \text{Tr}\, W(U)]. \qquad (4.2)$$

Here $U = e^{R\Phi}$ with $\Phi \sim \Phi + 2\pi i/R$. As in the topological theory, we take Φ purely imaginary so that U is unitary, and $H^c = \bar{H}$. Since H^c, H appear quadratically, we can immediately reduce this to a matrix model. Rescaling $H^c(t), H(t)$ and integrating them out, we get the partition function

$$Z \sim \frac{1}{\text{Vol}(U(N))} \int dU \, \frac{e^{-\frac{1}{g_s}\text{Tr}W(U)}}{\det(\partial_t + \Phi + m)}. \qquad (4.3)$$

Actually we only expect this for $k = 0$. For $k \neq 0$ there may be a modification. Just as in the previous section we can find the full answer by adding some hypermultiplets and taking their masses from minus to plus infinity. In general we get

$$Z \sim \frac{1}{\text{Vol}(U(N))} \int dU \, \det(U)^{-k} \frac{e^{-\frac{1}{g_s}\text{Tr}W(U)}}{\det(\partial_t + \Phi + m)}. \qquad (4.4)$$

In other words, when we take the $d^2\theta$ terms in the field theory action as the action for the matrix model, we should be careful to include a contribution $k\text{Tr}(\Phi)$ for the Chern-Simons term. This is not unreasonable[10] given that the field theory resolvent $\text{Tr}\, W_\alpha W^\alpha/(z - \Phi)$ (plus images if Φ is periodic) is replaced in the matrix model by $\text{Tr}\, 1/(z - \Phi)$.

It is tempting to state this in general: non-minimal kinetic terms of the form

$$\int dx d^2\theta \, \text{Tr}(f(\Phi)W_\alpha W^\alpha) \qquad (4.5)$$

should lead to a term $\text{Tr}(f(\Phi))$ in the matrix model action (without a factor of $1/g_s$). This can perhaps be shown in a similar way by integrating out matter with appropriate mass terms.

Now we proceed to apply some standard techniques.

---

[10] I would like to thank C. Vafa for this remark

4.1. *Deriving the matrix model curve.*

Diagonalising $U = \text{diag}(e^{R\phi_1}, \ldots, e^{R\phi_N})$ and using the formula

$$\det(\partial_t + x) \sim \det \frac{\sinh(\frac{1}{2}Rx)}{\frac{1}{2}R} \tag{4.6}$$

we obtain

$$
\begin{aligned}
Z &\sim \frac{1}{\text{Vol}} \int dU \, \frac{e^{-\frac{1}{g_s}\text{Tr}\,W(U) - k\,\text{Tr}\,R\Phi}}{\det \sinh \frac{1}{2}R(\Phi + m)} \\
&\sim \frac{1}{\text{Vol}} \int \prod d\phi_i \, \frac{\prod_{i \neq j} \sinh \frac{1}{2}R(\phi_i - \phi_j)}{\prod_i \sinh \frac{1}{2}R(\phi_i + m)} e^{-\frac{1}{g_s}\sum_i W(R\phi_i) - k\sum_i R\phi_i} \\
&\sim \frac{1}{\text{Vol}} \int \prod d\phi_i \, e^{-\frac{1}{g_s}\sum_i W(R\phi_i) + \sum_{i \neq j} \log \sinh \frac{1}{2}R(\phi_i - \phi_j) - \sum_i \log \sinh \frac{1}{2}R(m+\phi_i) - kR\phi_i}.
\end{aligned}
\tag{4.7}
$$

Notice that the Chern-Simons level manifests itself as an additional force term in the action, as expected from our discussion in section 3. The equations of motion are given by

$$-\frac{1}{g_s}W'(R\phi_i) + \sum_{j, j \neq i} \coth \frac{1}{2}R(\phi_i - \phi_j) - \frac{1}{2}\coth \frac{1}{2}R(m + \phi_i) - k = 0. \tag{4.8}$$

Introducing the resolvent

$$\omega(v) = -\frac{1}{N}\sum_i \coth \frac{1}{2}R(v - \phi_i) \tag{4.9}$$

we can bring expression (4.8) into the form

$$
\begin{aligned}
0 &= -\frac{1}{g_s N}f(v) + \frac{1}{g_s N}W'(Rv)\omega(v) + \frac{1}{2}[\omega^2(v) - \frac{2}{RN}\omega'(v) - 1] \\
&\quad - \frac{1}{2N^2}\sum_i \coth \frac{1}{2}R(m + \phi_i)\coth \frac{1}{2}R(v - \phi_i) + \frac{k}{N}\omega(v)
\end{aligned}
\tag{4.10}
$$

where the error term is given by

$$f(v) = \frac{1}{N}\sum_i \left(W'(R\phi_i) - W'(Rv)\right)\coth \frac{1}{2}R(v - \phi_i). \tag{4.11}$$

In the large N limit, with $S = g_s N$ fixed, this reduces to

$$-\frac{1}{g_s N}f(v) + \frac{1}{g_s N}W'(Rv)\omega(v) + \frac{1}{2}[\omega^2(v) - 1] = 0. \tag{4.12}$$

19

Defining

$$y = S\omega(v) + W'(Rv) \qquad (4.13)$$

this can equivalently be written as

$$y^2 = (W'(Rv))^2 + 2Sf(v) + S^2 = (W'(Rv))^2 + q(v). \qquad (4.14)$$

As mentioned before $q(v)$ must corresponds to a normalisable complex structure deformation. This means that, if b is a parameter in $q(v)$, then the integral of holomorphic form $\partial_b \lambda$ over a contour that runs to infinity must be convergent or at most logarithmically divergent. One of the differences of the five dimensional case compared to the four dimensional case is the existence of two asymptotic infinities (on each sheet). As a result of this we will see that there are two logarithmically divergent (1,0) forms as opposed to only one in the four dimensional case. However we will see only a particular combination of the two arises from the matrix model. Thus it appears that not all allowed complex structure deformations actually occur.

4.2. Relations between the coefficients.

A simple counting shows that there must be relations between the b_i. We consider two cases in turn[11]: $W(U)$ has only positive powers of U, or $W(U)$ contains both positive and negative powers of U.

<u>case i</u>. $W(U) = \sum_{i=0}^{n+1} g_i U^i$.

In this case W has n critical points at finite distance and one at $-\infty$ (from solving $W'(v) \sim x\partial W/\partial x = 0$). It is easy to see that in this case the quantum correction is of the form

$$q(x) = \sum_{i=0}^{n+1} b_i x^i. \qquad (4.15)$$

Moreover one can see that the constant term in $f(x)$ is zero because $\langle U\partial W/\partial U \rangle = 0$, and the leading term in $f(x)$ is $(n+1)g_{n+1}x^{n+1}$. Therefore there is a relation between b_0 and b_{n+1}:

$$\frac{b_{n+1}}{2(n+1)g_{n+1}} = \sqrt{b_0} = S. \qquad (4.16)$$

---

[11] In the first case W is a holomorphic function with a pole only at infinity. Therefore picking a 'contour' which encloses no poles in the evaluation of the partition function, such as integration over unitary matrices, gives identically zero. This should not be relevant however, since for field theory purposes we are not interested in defining the matrix model non-perturbatively. I would like to thank C. Roemelsberger for this remark.

Moreover this has the following implication for the periods of $\lambda_{mm} = y\,dx/x$:

$$\frac{1}{4\pi i} \oint_{v=\infty} \lambda = \frac{1}{4\pi i} \oint_{v=-\infty} \lambda = S/2. \qquad (4.17)$$

The contours around $\pm\infty$ are taken counterclockwise. Taking account of (4.16), this implies that the quantum correction has $n+1$ adjustable parameters. This is the right number: we have a total of $n+1$ critical points which get smeared out into cuts. Here we include the cut that degenerates to a double root at $-\infty$ as the quantum correction is turned off. Let us denote by A_i a contour which encircles the ith cut (we associate A_0 with the root at $-\infty$). Then we have $n+1$ independent variables S_i which are related to the b_i through

$$\frac{1}{4\pi i} \oint_{A_i} \lambda = S_i \quad i = 0, \ldots, n. \qquad (4.18)$$

<u>case ii</u>. $W(U) = \sum_{i=-1}^{n} g_i U^i$.

(we could allow for more powers of U^{-1}, but this doesn't add anything new). In this case the superpotential has $n+1$ critical points. However W has only n parameters to adjust the positions of the roots (the other is an overall scalar multiplying W). Thus the position of the last critical point is fixed in terms of the previous n.

The matrix model again implies a relation between the coefficients of the quantum correction. In this case it is easy to determine the highest and lowest terms in $f(x)$. We find that

$$\frac{b_n}{n g_n} = \frac{b_{-1}}{g_{-1}} = 2S \qquad (4.19)$$

so that again we have

$$\frac{1}{4\pi i} \oint_{v=\infty} \lambda = \frac{1}{4\pi i} \oint_{v=-\infty} \lambda = S/2. \qquad (4.20)$$

This gives the right number of parameters: there are $n+1$ independent b_i after taking into account (4.19), and there are $n+1$ variables S_i related to the b_i through

$$\frac{1}{4\pi i} \oint_{A_i} \lambda = S_i \quad i = 0, \ldots, n. \qquad (4.21)$$

Here we take the contours A_i, $i = 1, \ldots, n$ to encircle n independent critical points and the contour A_0 to encircle the remaining critical point.

4.3. Effective superpotential

The effective superpotential is given by the following formula

$$W_{\text{eff}} = \int h \wedge \lambda \tag{4.22}$$

where $\lambda = y\,dx/x$. We have discussed previously the correspondence between the field theory data and the periods of h using the string set-up. One should also be able to derive this directly from the matrix model as in [24], which was made more precise in appendix B of [22].

The A-periods of h were as follows:

$$\oint_{A_i} h = N_i \qquad \oint_{A_0} h = N_0 \qquad \oint_{-\infty} h = 0 \qquad \oint_{\infty} h = N$$
$$\oint_{-\tilde{\infty}} h = -(N - k_q) \qquad \oint_{\tilde{\infty}} h = -(k_q - 1) \qquad \oint_{A_{-\tilde{m}}} h = -1 \tag{4.23}$$

We use $\pm\infty$ for the two ends of the cylinder, and a tilde for points on the second sheet (which cannot be seen classically). Also $N = \sum N_i$, and $A_{-\tilde{m}}$ is a contour around the point $v = -m$ on the lower sheet.

We also need to specify the periods for the B-cycles. We take $\int_{C_i} h = 0$ for the compact cycles (where $C_i = B_i - B_0$) and[12]

$$\tfrac{1}{2}\int_{-\tilde{\infty}}^{-\infty} h + \tfrac{1}{2}\int_{+\tilde{\infty}}^{+\infty} h = \tau. \tag{4.24}$$

We believe this is the correct prescription by comparing with the four dimensional limit of the five dimensional gauge theory.

Now we expand the superpotential and obtain

$$
\begin{aligned}
W_{\text{eff}} &= \sum_{i=1}^{n} N_i \int_{C_i} \lambda + N \int_{-\tilde{\infty}}^{+\infty} \lambda + k \int_{+\tilde{\infty}}^{-\tilde{\infty}} \lambda + \int_{-\tilde{m}}^{+\tilde{\infty}} \lambda + 4\pi i \tau \sum_{i=0}^{n} S_i \\
&= \sum_{i=0}^{n} N_i \int_{B_i} \lambda + k \int_{+\tilde{\infty}}^{-\tilde{\infty}} \lambda + \int_{-\tilde{m}}^{+\tilde{\infty}} \lambda + 4\pi i \tau \sum_{i=0}^{n} S_i.
\end{aligned}
\tag{4.25}
$$

As mentioned there is in principle some more flexibility in the periods of h. We have chosen to distribute the N units of flux over the two infinities on the upper sheet in a

---

[12] The precise contour we take is not relevant since $\int_{C_i} h = 0$. In cases where it is ambiguous, contours written as $a \to b$ mean the following: a contour between two points on the same sheet is chosen to lie on that sheet, and a contour between two points on different sheets is chosen to cross the A_0 cycle but no other cycle A_i for $i > 0$.

particular way; we can shift the fluxes to $\oint_{-\infty} h = f$ and $\oint_{\infty} h = N - f$ provided we also shift the flux at $-\tilde{\infty}$ and $+\tilde{\infty}$ by $-f$ and $+f$ respectively. This should have no effect on the effective superpotential when written in terms of physical parameters only. Finally we can move the matter brane through a cut onto the classical sheet and take it to the ends of the cylinder, changing the fluxes at $\pm\infty$. This corresponds to a Higgsed phase of the field theory [25].

4.4. Minimisation of the superpotential, and factorisation of the curve.

In the remainder we would like to apply the strategy employed in [26] to extract the Seiberg-Witten curve. The idea is to pick an arbitrary point on the Coulomb branch, i.e. choose N points on the cylinder, and consider a superpotential W for which these N points are all critical points.[13] Then in the matrix model we put one eigenvalue at each critical point. After minimising the effective superpotential, we can scale the superpotential to zero. The $U(1)^N$ coupling constants are unchanged under the rescaling of the superpotential and so agree with the coupling constants of the $\mathcal{N} = 2$ theory obtained from the Seiberg-Witten curve.

In the four dimensional case it was possible to choose $W'(x) = \det(x - \Phi)$ (for pure gauge theory), so that the matrix model curve reduces to the Seiberg-Witten curve up to a rescaling. For the five-dimensional case however, for the superpotentials considered above we have seen that we cannot choose all the critical points independently.

We will now try to recover the Seiberg-Witten curve by choosing a classical superpotential W which has N independent critical points and an additional dependent critical point (corresponding to the A_0 cycle for cases i and ii above).

Thus we take $N_0 = 0, N_{i>0} = 1$ and minimise with respect to the S_i. We have noted before that there are $n + 1$ independent variables S_i and $n + 1$ independent coefficients b_i in $q(x) = \sum x^i b_i$. So provided the Jacobian for the change of variables is non-singular, we can equivalently minimise with respect to the b_i. We expect that when we solve and substitute back, the cut where we did not put any eigenvalues will degenerate to a double point, so that the matrix model curve factorises as

$$y^2 = H(x)^2 F(x). \tag{4.26}$$

---

[13]   Actually we will see that in the five dimensional case the relation between the point on the Coulomb branch and the critical points of W is a little more complicated.

The details for cases i and ii introduced above will diverge due to the fact that the relation between the b_i's is slightly different. We will treat case ii here.

For case ii, $\partial_{b_i}\lambda$ for $i = 0, \ldots, N-1$ span the space of holomorphic 1-forms on the matrix model curve, which has genus N. For $i = -1, N$ we find that $\partial_{b_i}\lambda$ is logarithmically divergent at one of the ends of the cylinder, but we can not consider these two independently because of the relation between b_{-1} and b_N (4.19).

Differentiating W_{eff} with respect to b_j for $j = 0, \ldots n-1$ we find that

$$\sum_{i=1}^{n} N_i \int_{C_i} \partial_{b_j}\lambda = 0 \tag{4.27}$$

and

$$N \int_{-\infty}^{+\infty} \partial_{b_j}\lambda + k_q \int_{+\tilde{\infty}}^{-\tilde{\infty}} \partial_{b_j}\lambda + \int_{-\tilde{m}}^{+\tilde{\infty}} \partial_{b_j}\lambda = 0. \tag{4.28}$$

Invoking Abel's theorem, this implies there must exist a meromorphic function on the matrix model curve with a pole of order N at ∞, a zero of order $k_q - 1$ at $+\tilde{\infty}$ a zero of order $N - k_q$ at $-\tilde{\infty}$ and a zero of order 1 at $-\tilde{m}$, and regular elsewhere. Such a function is of the form

$$z \sim P_N(x) + \sqrt{P_N(x) - \gamma x^{N-k_q}(x - e^{-Rm})}. \tag{4.29}$$

with $P_N(x) = \sum_i p_i x^i$ a polynomial in x of degree N. We take $p_N = 1$ and $p_0 \neq 0$. Because of the square root, such a function can only exist on a curve of genus $N-1$. Then the matrix model curve must factorise as

$$y^2 = (x\,\partial_x W)^2 + q(x) = g^2(x^{-1} - a)^2(P_N^2(x) - \gamma x^{N-k_q}(x - e^{-Rm})). \tag{4.30}$$

The parameter a can be used to ensure that the condition (4.19) for b_{-1} and b_N is satisfied. The constant g is an overall scale in the superpotential. We can take $g \to 0$ to turn off the superpotential without changing the positions of the critical points. The last available parameter γ is fixed by solving

$$\partial_S W_{\text{eff}} \equiv 2(N g_N \partial_{b_N} + g_{-1}\partial_{b_{-1}})W_{\text{eff}} = 0. \tag{4.31}$$

Clearly the last factor in (4.30) looks a lot like the expected Seiberg-Witten curve. It needs to be shown that for a given $P_N(x)$ one can construct a superpotential $W(x)$ with the property that after minimisation of the effective superpotential $W_{\text{eff}}(S)$ the matrix

24

model curve factorises as in (4.30) with the expected polynomial $P_N(x)$. To simplify the algebra a bit, let us assume there is no matter multiplet. Then we have

$$y^2 = (x\,\partial_x W)^2 + q(x) = g^2(x^{-1} - a)^2(P_N^2(x) - \gamma x^{N-k}). \tag{4.32}$$

As an Ansatz for the relation between W' and P_N we write

$$x\,\partial_x W = g(x^{-1} - a)P_N(x) + gc + gdx^{-1}. \tag{4.33}$$

We can get an equation for the constant c by requiring that the right-hand-side has no constant term (otherwise we would get a $\log(x)$ term in W, which does not respect the periodic identifications imposed by gauge invariance). So

$$c = -p_1 + ap_0. \tag{4.34}$$

The constant d is fixed by requiring that there is no x^{-2} term in $q(x)$. Expanding $x\,\partial_x W$ in (4.32) we find that

$$q(x) = -2g^2(x^{-1} - a)(c + dx^{-1})P_N(x) - g^2(c + dx^{-1})^2 - g^2(x^{-1} - a)^2\gamma x^{N-k}. \tag{4.35}$$

One can check it is always possible to satisfy (4.19) for any of the allowed values $k = 0, \ldots, N$. So for every allowed $P_N(x)$ there exists a superpotential $W(x)$ for which (4.32) is realised as a critical point of the effective superpotential $W_{\text{eff}}(S)$.

For $k = 0$ we find $d = 0, c = \gamma p_1/(\gamma + 4p_0)$. For 'generic' k ($k = 1, \ldots, N - 1$) we seem to find two branches. If $d = 0$ then also $c = 0$, which looks like the vacuum we are after. But if $d = -2p_0$ then c can be anything. The latter branch seems puzzling because we have used $N - 1$ equations but still have two undetermined variables (c and γ), so the glueball superpotential appears to have a flat direction. Moreover on this branch we fail to have the property that $c, d \to 0$ as $\gamma \to 0$ so that the roots of P_N do not 'classically' coincide with critical points of W. We suspect this branch can be eliminated by a more careful analysis.

For $k = N - 1$ we find $d = 0$ and $c = \gamma/(4p_0)$. Finally for $k = N$ we find two solutions, depending on the sign in $d = -p_0 \pm \sqrt{p_0^2 + \gamma}$. Requiring $c, d \to 0$ as $\gamma \to 0$ singles out $d = -p_0 + \sqrt{p_0^2 + \gamma}$ as the expected solution, and the other solution should presumably be eliminated.

Notice that for several values of k we found that the roots of $P_N(x)$ are slightly displaced with respect to the critical points of $W(x)$. This should presumably be interpreted as arising from instanton corrections to the classical expectation values $\langle \text{Tr}(U_{\text{cl}}^m) \rangle$.

Thus it appears we have recovered the Seiberg-Witten curve from the matrix model curve (4.32). However this is meaningless if we don't recover the correct Seiberg-Witten differential, which amongst others gives the $U(1)$ coupling constants and the identification of the coefficients of P_N with Coulomb branch parameters in the $\mathcal{N} = 2$ theory.

4.5. Seiberg-Witten differential

Recall that in the string theory embedding the D-branes are replaced in the large N limit by fluxes. Therefore in the effective theory the gauge theory density of eigenvalues of Φ is given by h, which can be used to compute quantum expectation values. On the other hand the Seiberg-Witten differential does the same thing, so $\lambda_{SW} = \log(x)\,h$. This agrees with the Seiberg-Witten differential obtained from geometric engineering provided $h = -\frac{1}{2\pi i}d\log(z)$. We proceed to check that this is indeed the case.

The holomorphic 1-form $-\frac{1}{2\pi i}d\log(z)$ has the same periods over A-cycles as h. Both also have zero periods over the compact B-cycles $C_i = B_i - B_0$. We only need to check non-compact B-cycles.

From the variation with respect to S (4.31) we find:

$$\sum_{i=1}^{n} N_i \int_{C_i} \omega_S + N \int_{-\tilde{\infty}}^{+\infty} \omega_S + k_q \int_{+\tilde{\infty}}^{-\tilde{\infty}} \omega_S + \int_{-\tilde{m}}^{+\tilde{\infty}} \omega_S + 4\pi i\tau = 0. \tag{4.36}$$

where we put

$$\omega_S = 2(Ng_N\partial_{b_N} + g_{-1}\partial_{b_{-1}})\lambda, \tag{4.37}$$

and τ is given by (4.24). Since $\sum_{i=1}^{n} N_i \int_{C_i} \partial_{b_j}\lambda = 0$ for $j = 0, \ldots, n-1$ and $\partial_S\lambda$ is holomorphic away from the asymptotic regions, we get $\sum_{i=1}^{n} N_i \int_{C_i} \omega_S = 0$. Next we introduce the cut-offs $x \sim (R\Lambda_+)^2, x \sim (R\tilde{\Lambda}_+)^2$ for $+\infty, +\tilde{\infty}$ and $x \sim 1/(R\tilde{\Lambda}_-)^2$ for $-\tilde{\infty}$. As mentioned before we can interpret these as masses of additional hypermultiplets, eg. $R\Lambda_+ \sim e^{RM_+/2}$. Then

$$2\pi i\tau = -\log((R\Lambda_+)^N (R\tilde{\Lambda}_+)^{k-1}(R\tilde{\Lambda}_-)^{N-k}) + C + \mathcal{O}(1/R\Lambda_+, 1/R\tilde{\Lambda}_\pm) \tag{4.38}$$

for some constant C independent of the lambdas. We have to match this with (4.24). Unfortunately we have not been able to compute the subleading terms, so we will only check that the leading divergent pieces agree. Using $h = -\frac{1}{2\pi i}d\log(z)$ with z given in (4.29), we get

$$2\pi i\tau = \frac{1}{2}\log\left(\frac{\gamma^2 e^{-Rm}}{(-16p_0^2)(R\Lambda_+)^{2N}(R\tilde{\Lambda}_+)^{2k-2}(R\tilde{\Lambda}_-)^{2N-2k}}\right) \tag{4.39}$$

which indeed has the same divergent behaviour.

To get agreement with (2.26) we want to identify $\gamma \sim 4(R\Lambda)^{2N-1}e^{Rm/2}$. Then we identify $(R\Lambda_+)^N (R\tilde{\Lambda}_+)^{k-1}(R\tilde{\Lambda}_-)^{N-k}$ with the gauge theory cut-off $(R\Lambda_0)^{2N-1}$ (the factor

of $-p_0^2$ presumably gets taken care of by a counterterm $\pi i S + \log(p_0^2)S$ to the superpotential; see [22] for similar terms in four dimensions).

References

[1] R. Dijkgraaf and C. Vafa, "Matrix models, topological strings, and supersymmetric gauge theories," Nucl. Phys. B **644**, 3 (2002) [arXiv:hep-th/0206255].
R. Dijkgraaf and C. Vafa, "On geometry and matrix models," Nucl. Phys. B **644**, 21 (2002) [arXiv:hep-th/0207106].
R. Dijkgraaf and C. Vafa, "A perturbative window into non-perturbative physics," arXiv:hep-th/0208048.

[2] R. Dijkgraaf and C. Vafa, "N = 1 supersymmetry, deconstruction, and bosonic gauge theories," arXiv:hep-th/0302011.

[3] N. Nekrasov, "Five dimensional gauge theories and relativistic integrable systems," Nucl. Phys. B **531**, 323 (1998) [arXiv:hep-th/9609219].

[4] B. Kol, "5d field theories and M theory," JHEP **9911**, 026 (1999) [arXiv:hep-th/9705031].

[5] A. Brandhuber, N. Itzhaki, J. Sonnenschein, S. Theisen and S. Yankielowicz, "On the M-theory approach to (compactified) 5D field theories," Phys. Lett. B **415**, 127 (1997) [arXiv:hep-th/9709010].

[6] A. E. Lawrence and N. Nekrasov, "Instanton sums and five-dimensional gauge theories," Nucl. Phys. B **513**, 239 (1998) [arXiv:hep-th/9706025].

[7] H. Kanno and Y. Ohta, "Picard-Fuchs equation and prepotential of five dimensional SUSY gauge theory compactified on a circle," Nucl. Phys. B **530**, 73 (1998) [arXiv:hep-th/9801036].
T. Eguchi and H. Kanno, "Five-dimensional gauge theories and local mirror symmetry," Nucl. Phys. B **586**, 331 (2000) [arXiv:hep-th/0005008].

[8] A. Iqbal and V. S. Kaplunovsky, "Quantum deconstruction of a 5D SYM and its moduli space," arXiv:hep-th/0212098.
A. Iqbal and V. S. Kaplunovsky, To Appear.

[9] N. Nekrasov and A. Okounkov, "Seiberg-Witten theory and random partitions," arXiv:hep-th/0306238.

[10] T. J. Hollowood, "Five-dimensional gauge theories and quantum mechanical matrix models," JHEP **0303**, 039 (2003) [arXiv:hep-th/0302165].

[11] T. J. Hollowood, A. Iqbal and C. Vafa, "Matrix models, geometric engineering and elliptic genera," arXiv:hep-th/0310272.

[12] S. Ganguli, O. J. Ganor and J. A. Gill, "Twisted six dimensional gauge theories on tori, matrix models, and integrable systems," arXiv:hep-th/0311042.

[13] S. Katz, A. Klemm and C. Vafa, "Geometric engineering of quantum field theories," Nucl. Phys. B **497**, 173 (1997) [arXiv:hep-th/9609239].
S. Katz, P. Mayr and C. Vafa, "Mirror symmetry and exact solution of 4D N = 2 gauge theories. I," Adv. Theor. Math. Phys. **1**, 53 (1998) [arXiv:hep-th/9706110].

[14] K. A. Intriligator, D. R. Morrison and N. Seiberg, "Five-dimensional supersymmetric gauge theories and degenerations of Calabi-Yau spaces," Nucl. Phys. B **497**, 56 (1997) [arXiv:hep-th/9702198].

[15] O. Aharony, A. Hanany and B. Kol, "Webs of (p,q) 5-branes, five dimensional field theories and grid diagrams," JHEP **9801**, 002 (1998) [arXiv:hep-th/9710116].

[16] N. C. Leung and C. Vafa, "Branes and toric geometry," Adv. Theor. Math. Phys. **2**, 91 (1998) [arXiv:hep-th/9711013].

[17] E. Witten, "Phase Transitions In M-Theory And F-Theory," Nucl. Phys. B **471**, 195 (1996) [arXiv:hep-th/9603150].

[18] H. Ooguri and C. Vafa, "Knot invariants and topological strings," Nucl. Phys. B **577**, 419 (2000) [arXiv:hep-th/9912123].

[19] C. Vafa, "Superstrings and topological strings at large N," J. Math. Phys. **42**, 2798 (2001) [arXiv:hep-th/0008142].

[20] M. Aganagic, K. Intriligator, C. Vafa and N. P. Warner, "The glueball superpotential," arXiv:hep-th/0304271.
K. Intriligator, P. Kraus, A. V. Ryzhov, M. Shigemori and C. Vafa, "On low rank classical groups in string theory, gauge theory and matrix models," arXiv:hep-th/0311181.

[21] K. D. Kennaway and N. P. Warner, "Effective superpotentials, geometry and integrable systems," arXiv:hep-th/0312077.

[22] F. Cachazo, N. Seiberg and E. Witten, "Chiral Rings and Phases of Supersymmetric Gauge Theories," JHEP **0304**, 018 (2003) [arXiv:hep-th/0303207].

[23] N. Arkani-Hamed, T. Gregoire and J. Wacker, "Higher dimensional supersymmetry in 4D superspace," JHEP **0203**, 055 (2002) [arXiv:hep-th/0101233].

[24] S. G. Naculich, H. J. Schnitzer and N. Wyllard, "Matrix model approach to the N = 2 U(N) gauge theory with matter in the fundamental representation," JHEP **0301**, 015 (2003) [arXiv:hep-th/0211254].
S. G. Naculich, H. J. Schnitzer and N. Wyllard, "The N = 2 U(N) gauge theory prepotential and periods from a perturbative matrix model calculation," Nucl. Phys. B **651**, 106 (2003) [arXiv:hep-th/0211123].

[25] F. Cachazo, N. Seiberg and E. Witten, "Phases of N = 1 supersymmetric gauge theories and matrices," JHEP **0302**, 042 (2003) [arXiv:hep-th/0301006].

[26] F. Cachazo and C. Vafa, "N = 1 and N = 2 geometry from fluxes," arXiv:hep-th/0206017.